

Varied diversification patterns and distinct demographic trajectories in Ethiopian montane forest bird (Aves: Passeriformes) populations separated by the Great Rift Valley

Joseph Manthey¹, Yann Bourgeois², Yonas Meheretu³, and Stephane Boissinot¹

¹New York University - Abu Dhabi

²NYUAD

³Mekelle University

December 14, 2021

Abstract

Taxon-specific characteristics and extrinsic climatic and geological forces may both shape population differentiation and speciation. In geographically and taxonomically focused investigations, differentiation may occur synchronously as species respond to the same external conditions. Conversely, when evolution is investigated in taxa with largely varying traits, population differentiation and speciation is complex and shaped by interactions of Earth's template and species-specific traits. As such, it is important to characterize evolutionary histories broadly across the tree of life, especially in geographic regions that are exceptionally diverse and under pressures from human activities such as in biodiversity hotspots. Here, using whole-genome sequencing data, we characterize genomic variation in populations of six Ethiopian Highlands forest bird species separated by a lowland biogeographic barrier, the Great Rift Valley (GRV). In all six species, populations on either side of the GRV exhibited significant but varying levels of genetic differentiation. Species' dispersal ability was negatively correlated with levels of population differentiation. Isolation with migration models indicated varied patterns of population differentiation and connectivity among populations of the focal species. We found that demographic histories—estimated for each individual—varied by both species and population but were consistent between individuals of the same species and sampling region. We found that genomic diversity varied by half an order of magnitude across species, and that this variation could largely be explained by the harmonic mean of effective population size over the past 200,000 years. Overall, we found that even in highly dispersive species like birds, the GRV acts as a substantial biogeographic barrier.

INTRODUCTION

Population differentiation and speciation is a complex process that depends on species-specific factors as well as extrinsic landscape processes. An ongoing question in evolutionary biology is whether taxa differentiate in communities in response to extrinsic forces—for example, geographic or climatic change—or in species-specific manners based on their intrinsic characteristics such as physiological thermal tolerance or dispersal ability. In speciation scenarios dominated by geologic or climatic change, entire communities may differentiate in pulses synchronous with the changes in the Earth's template (e.g., Barber & Klicka, 2010; Musher et al., 2019; O'Connell et al., 2018). In contrast, when communities have species that vary in dispersal ability, niche breadth, or population history, diversification may be asynchronous across taxa based on those species' intrinsic characteristics (e.g., Naka & Brumfield, 2018; Oswald et al., 2017; Papadopoulou et al., 2009).

In well-studied geographic regions, it has become clear that with increasing taxonomic scope, a combination of both synchronous and asynchronous diversification arises that depends on the taxonomic and temporal breadth of study (Shafer et al., 2010; Smith et al., 2014). These results demonstrate the need for sampling multiple taxonomic groups, with a range of intrinsic characteristics and histories, to best understand the

drivers of diversification in different geographic contexts. This work is arguably most needed in biodiversity hotspots, the regions of the world that are richest in diversity while simultaneously most at risk due to human activities (Myers et al., 2000; Zachos & Habel, 2011). Not only are these regions some of the richest in biodiversity on our planet, but undescribed biodiversity—especially species and genetic diversity—is often concentrated in these regions (Hamilton et al., 2010; Miraldo et al., 2016; Mora et al., 2011).

The Horn of Africa biodiversity hotspot is rich in species (Friis et al., 2001; Largen & Spawls, 2010; Yalden & Largen, 1992), landscape perturbation (Dessie & Kleman, 2007; Zeleke & Hurni, 2001), and elevational heterogeneity; a majority of the topographic complexity is found in Ethiopia, with elevations ranging from more than a hundred meters below sea level in the Danakil Depression to greater than 4500 meters above sea level in the Simien Mountains (Fig. 1). The Ethiopian Highlands are a largely continuous region of tropical high elevation habitats with major lowland biogeographic barriers including the Great Rift Valley (GRV) and the Blue Nile Valley (Fig. 1). Both the Blue Nile Valley and the GRV are part of the large East African rift system (Frisch, Meschede, & Blakey, 2010). These lowland biogeographic barriers have shaped geographic population structure in a variety of montane species, including mammals, frogs, and plants (Belay & Mori, 2006; Bryja et al., 2018; Evans et al., 2011; Freilich et al., 2016; Gottelli et al., 2004; Kebede et al., 2007; Manthey et al., 2017; Reyes-Velasco et al., 2018; Reyes-Velasco et al., 2018; Silvestrini et al., 2007; P. J. Taylor et al., 2011).

Birds are often assumed to be good dispersers, making them a good focal taxon to identify whether the GRV is a significant biogeographic barrier for relatively highly dispersive species. Despite the general assumptions of bird dispersal ability due to flight, tropical birds—especially those that are non-migratory—may not always disperse long distances. For example, rivers have been shown to be long-term dispersal barriers in Amazonian birds (Naka & Brumfield, 2018). Additionally, isolated sky islands in other regions of the East African rift system have promoted diversification in some avian taxa, suggesting at least some montane birds have limited dispersal across lowland biogeographic barriers (Habel et al., 2015). These patterns suggest that even in birds, species-specific diversification patterns may exist due to intrinsic characteristics of each species such as dispersal ability.

What remains lacking are studies of comparative population structure in Ethiopian montane birds; to date, there have been no studies investigating phylogeographic or population genetic patterns in this diverse community. To help fill this gap, we used a comparative framework to study the effects of the GRV on population genetic differentiation in montane forest birds of the Ethiopian Highlands. We included six bird species in this study: *Cossypha semirufa* (Rüppell’s Robin-chat), *Crithagra tristriata* (Brown-rumped Seedeater), *Melaenornis chocolatinus* (Abyssinian Slaty Flycatcher), *Sylvia galinieri* (Abyssinian Catbird), *Turdus abyssinicus* (Abyssinian Thrush), and *Zosterops poliogastrus* (Ethiopian White-eye) (Table 1). We chose these species for several reasons, as they are all (1) highland forest species in Ethiopia, (2) relatively common where found locally, and (3) endemic to Ethiopia or the Horn of Africa region. These six bird species are associated with various types of forests and woodlands, including forest edge, and they can often be found in the same communities; indeed, we observed and captured all species for this study in the same general sampling areas (Fig. 1). However, some of these species have different habitat associations and minimum elevational affinities (Fig. 2). For example, the Abyssinian Catbird and Rüppell’s Robin-chat often prefer *Juniperus* and *Podocarpus* forests, the Abyssinian Thrush and Brown-rumped Seedeater are occasionally found in highland scrub habitat, and the Abyssinian Slaty Flycatcher is often found near woodland streams (Clement, 2020; Collar, 2020; Collar & Robson, 2021; del Hoyo, Collar, & Kirwan, 2020; B. Taylor, 2020; van Balen et al., 2020). In addition to differential habitat preferences, the species studied here have different wing shapes as measured by the hand-wing index (HWI; Table 1) (Claramunt et al., 2011; Kipp, 1959). Because the HWI is positively correlated with dispersal ability in birds, we may expect species with higher HWI, such as the Abyssinian Thrush and Ethiopian White-eye, to have maintained relatively higher population connectivity relative to other species studied here.

We used genome resequencing data to estimate genomic diversity and differentiation, timing of diversification, and demographic histories for these six bird species on either side of the GRV. We then tested whether species-

specific characteristics, including dispersal ability and demographic history, could explain the comparative patterns of population genomic diversity and differentiation.

METHODS

Fieldwork and sampling

We went on two expeditions in 2016 and 2017 to acquire genetic samples from populations of several bird species on both sides of the Great Rift Valley (GRV) (Fig. 1). We mist-netted birds and sampled: (1) blood, (2) wing measurements, and (3) photographs of each individual. All fieldwork was done in accordance with NYU IACUC accepted procedures (IACUC Protocol Number 15-00002A1 to SB) and permissions from the Ethiopian Wildlife Conservation Authority and the Oromia Forest and Wildlife Enterprise. For this study, we included six species from which we captured individuals on both sides of the GRV (Table 1). Of note, we were limited to three samples per species per region because of permitting regulations.

We sampled a total of [?] 20 μ L of blood from each individual for genomic resources. As a proxy for dispersal ability, we took measurements of the (1) wrist to tip of longest primary and (2) length of first secondary feather to calculate the hand-wing index (HWI) for each individual (Claramunt et al., 2011; Kipp, 1959). Lastly, we took profile photographs of each bird. Information of genetic samples, bird photos, locality information, and habitat photos are accessioned at the University of Kansas Biodiversity Institute, with much of the information presented in Table S1.

Population genomic resequencing, data processing, and genotyping

Sequencing. We resequenced a total of 30 individuals from two populations each of six montane bird species from the Ethiopian Highlands (Fig. 1, Table 1, Table S1). We extracted genomic DNA from blood samples using QIAGEN (Hilden, Germany) DNeasy blood and tissue kits following manufacturer guidelines. Genomic DNA extractions were sent to the Oklahoma Medical Research Foundation (OMRF) Clinical Genomics Center for standard Illumina shotgun sequencing library creation and subsequent sequencing on either an Illumina HiSeq3500 or NovaSeq6000. Samples were multiplexed on flow cells with other libraries from unrelated projects, with the goal to obtain \sim 8-25x mean genomic sequencing coverage. We were able to aim for similar sequencing coverage across samples because avian genomes tend to have very similar sizes (\sim 1.2 Gbp).

Filtering and Genotyping. We used the `bbduk.sh` script of the `bbmap` package (Bushnell, 2014) to trim adapters and quality filter raw sequencing data. We used BWA (Li & Durbin, 2009) with the `BWA-MEM` command to align filtered reads to a reference genome. Here, we used the Zebra Finch (*Taeniopygia guttata*) genome (Warren et al., 2010) as a reference for aligning reads and genotyping. This is possible as all birds in this study are in the same order Passeriformes, and most of the genome is similar enough to the species studied here so as to align reads (Fig. S1). In addition, because synteny is well conserved in birds (Derjushcheva et al., 2004; Griffin et al., 2008), analyses relying on haplotype structure should not be strongly biased. Despite being able to align most regions of the genome (\sim 76% to 86% in each individual; Table S1), some regions may not be aligned between our focal species and the reference genome due to (1) poorly assembled regions of the reference, (2) regions evolving quickly in the Zebra Finch or our focal species studied here, or (3) hard to align regions such as non-unique regions of the genome (e.g, repetitive elements, etc.).

We used `samtools v1.4.1` (Li et al., 2009) to convert the BWA output SAM file to BAM format, and lastly cleaned, sorted, added read groups to, and removed duplicates from each BAM file using the Genome Analysis Toolkit (GATK) v4.1.0.0 (McKenna et al., 2010). We then used GATK's functions `HaplotypeCaller` and `GenotypeGVCFs` to group genotype each species' individuals together for both variant and invariant sites. Here, we limited analyses to chromosomes with a 1 Mbp minimum size. We measured the distribution of sequencing coverage using the `samtools 'depth'` command (Fig. S1).

We used `VCFtools v0.1.14` (Danecek et al., 2011) to initially filter all variant and invariant site calls using the following restrictions: (1) minimum site quality of 20, (2) minimum genotype quality of 20, (3) minimum depth of coverage of 6, (4) maximum mean depth of coverage of 50, and (5) indel removal.

Mitochondrial Genomes. We assembled complete mitochondrial genomes for each individual and uploaded them to GenBank as a resource for the research community. Briefly, we extracted potential mitochondrial reads from the raw FASTQ data using the `bbsplit.sh` script from the BBSplit package (Bushnell, 2014) and several songbird mitogenomes as references. We used these putative mtDNA reads in Geneious (BioMatters Ltd.) to assemble mitogenomes. We then used the Geneious “Live Annotate” feature with the reference songbird mitogenomes to annotate the new assemblies. All assemblies are accessioned on NCBI’s GenBank: MT017889- MT017917. To visualize mtDNA structure, we created a median joining network of the NADH Dehydrogenase Subunit 2 (ND2) gene for each species using PopART (Bandelt, Forster, & Röhl, 1999; Leigh & Bryant, 2015).

Population genomics

Diversity and differentiation. We estimated summary statistics in 500 kbp sliding, non-overlapping windows along the genome. We chose this window size to ensure we obtained a large number of variable sites per window, given the small sample sizes per taxon. Differentiation and divergence statistics (F_{ST} and D_{XY}) were calculated separately for each species, and only sites with no missing genotypes within a species were used. Heterozygosity was measured for all genotyped sites passing filters per individual. We used custom R scripts to measure observed heterozygosity for each individual and two measures of genetic differentiation (F_{ST} and D_{XY}) between populations on either side of the GRV. We used the Reich and colleagues (2009) estimator of F_{ST} because this has been shown to be an unbiased F_{ST} estimator when using low sample sizes (e.g., two or more chromosomes per population) and high numbers of genetic markers (Willing, Dreyer, & Van Oosterhout, 2012). We calculated both F_{ST} and D_{XY} to measure how they covary across the genome, as well as have measures of genetic differentiation that both do (F_{ST}) and do not (D_{XY}) rely on within population genetic diversity. If the genomes of differentiating populations are largely under the influence of linked selection, we would expect a genome-wide negative relationship between F_{ST} and D_{XY} . In contrast, if differentiating populations exhibit gene flow across most of their genomes, with restricted gene flow in only small genomic regions (e.g., genomic islands of differentiation), we would expect a genome-wide positive relationship between F_{ST} and D_{XY} .

Population genetic structure. We estimated population genetic structure with an a priori assumption of two populations using the program STRUCTURE (Pritchard, Stephens, & Donnelly, 2000). Here, our goal was to assess whether we could identify distinct population clusters with individuals segregating on either side of the GRV despite largely different levels of genetic differentiation between populations in different species (see RESULTS; Table 1). We limited SNPs for each species to those that were separated by a minimum of 20 kbp to reduce effects of linkage as well as reduce computation time. This resulted in 42,542 to 49,443 SNPs per species. We initially ran STRUCTURE to infer the lambda parameter while estimating the likelihood of one population ($k = 1$) (Pritchard et al., 2000). We then used the inferred value of lambda for subsequent analyses, where we performed five replicates of likelihood estimation for two genetic clusters. We assumed correlated allele frequencies, an admixture model, and performed analyses for a burn-in of 100,000 steps and a subsequent 500,000 MCMC iterations.

Phylogenomics. We estimated phylogenies for all the individuals combined to (1) identify phylogenetic splits between populations separated by the GRV and (2) analyze trait data (e.g., HWI, genetic diversity) in the context of phylogenetic independent contrasts (PICs). Here, we estimated gene trees from non-overlapping 100 kbp alignments using RAxML v8.2.12 (Stamatakis, 2014) with the GTRCAT model of sequence evolution. For sites to be included, we required 80% of all individuals to be genotyped at that site and included only invariant and biallelic sites. For an alignment’s inclusion, it needed at least 10 kbp retained sites. We created a consensus tree from the gene trees using the `sumtrees.py` script, part of the DendroPy Python package (Sukumaran & Holder, 2010), for use in calculating PICs. PICs are used to estimate correlations among traits while accounting for evolutionary history of the samples (i.e., because sampled lineages are not independent from one another). The PIC method assumes Brownian motion of trait evolution, and transforms the sampled data (i.e., at the tips of the phylogeny) into statistically independent contrasts that may be used in regressions (Felsenstein, 1985). We estimated PICs in the R package `ape` (Paradis, Claude,

& Strimmer, 2004). We estimated a species tree using ASTRAL III v 5.6.3 (Zhang et al., 2018), using the quartet frequencies as a measure of local support (Sayyari & Mirarab, 2016).

Demography. We used the program MSMC2 v1.1.0 (Schiffels & Durbin, 2014) to estimate demographic history for each individual. MSMC uses information about the spatial arrangement of variant sites within an individual’s two chromosomes using a modified version of the Pairwise Sequentially Markovian Coalescent (PSMC). We masked regions of the genome that were not genotyped, as MSMC would otherwise mistake these regions as runs of homozygosity. Of note, the MSMC method is accurate in panmictic populations, but the presence of population structure or changes in connectivity between populations through time may mimic changes in population sizes (Chikhi et al., 2018; Mazet et al., 2016). Therefore, some caution should be used when interpreting the demographic histories. We ran MSMC for each individual allowing up to 20 iterations (default setting), and up to 23 inferred distinct time segments. To determine the number of time segments to use, we ran preliminary MSMC runs with the default number of time segments ($n = 28$), and decreased that number in subsequent runs until spurious results and model overfitting were diminished. For example, we merged some recent time segments to prevent overfitting in the most recent (e.g., past 50 kya) and very old times, as these would sometimes have spurious jumps in population sizes to unrealistically high values of N_E (e.g., many millions). To assess how signal may vary using different parts of the genome, we performed ten bootstrap replicates for each individual, bootstrapping 1 Mbp segments of the genome. Here, we chose to run MSMC for each individual rather than aggregating data from individuals for several reasons: (1) we have low to moderate sequencing coverage for each of the samples, and masking low coverage regions across multiple individuals drastically decreases the amount of the genome sampled, (2) to reduce the impacts of uneven sample sizes per population and/or species, and (3) with demographic results for each individual we can assess consistency across individuals within species and populations.

The output of MSMC is presented in terms of a species’ generation times and mutation rate. As such, the observed demographic patterns are reliant on accurate estimates of both these parameters. In the species studied here, there are no published estimates of generation times. As such we used a proxy used previously in the literature, where we use a conservative generation time of double the age of sexual maturity in closely-related species (Nadachowska-Brzyska et al., 2015). We estimated the age of sexual maturity for each species by finding closely-related species (same genus or closely-related genus) using the Animal Aging and Longevity Database (Tacutu et al., 2017) (available here: genomics.senescence.info/species). We discuss caveats associated with assumed generation times in the RESULTS section.

Because no estimates for rates of molecular evolution exist for any of the species in this study, we estimated substitution rates for each species included here. To do this, we extracted coding regions of the Zebra Finch genome that were genotyped in at least one individual from each species included in this study using the BEDtools v2.27.1 (Quinlan & Hall, 2010) *intersect* command, with the Zebra Finch CDS regions and the GATK vcf outputs as input for BEDtools. We filtered any CDS regions that included the same start codon to remove overlapping alignments. Any null alleles or missing data in alignments were replaced with Ns and only alignments that included a minimum of 95% of the coding region were included for further analysis. Any codons including Ns were removed from the alignments. Lastly, we extracted all four-fold degenerate sites and concatenated them into a single alignment using the R packages Biostrings, seqinr, and rphast (Charif & Lobry, 2007; Hubisz, Pollard, & Siepel, 2010; Pagès et al., 2017). The final alignment included a total of 496,080 sites. From this alignment, we estimated a model of sequence evolution using jModelTest (Darriba et al., 2012) and chose a best-fitting model with the Akaike Information Criterion (best model = GTR + I + G). Next, we estimated a phylogeny using maximum likelihood with the program PhyML (Guindon et al., 2010).

We used the branch lengths of the resulting phylogeny to estimate substitution rates for each taxon. Here, we used split times for each of the respective genera from a fossil-calibrated phylogenetic tree of Passeriformes (Oliveros et al., 2019), where we could put the amount of evolution on each branch in our four-fold degenerate sites phylogeny in the context of time. All of the rates calculated here ranged between 2.06×10^{-9} and 2.30×10^{-9} substitutions / site / year (Table 1) and are similar to other Passeriformes species such as the Zebra

Finch (2.21×10^{-9} substitutions / site / year) (Nam et al., 2010). These mutation rate estimates assume that birds in the same genera do not have greatly differing terminal branch lengths, as this would erroneously either increase or decrease mutation rates. Regardless, these rates are similar to those estimated in other Passeriformes species, and even with some error, we expect these estimates to be broadly appropriate for general interpretations of demographic histories.

Population divergence and connectivity through time.

To estimate population divergence and connectivity through time, we used the program MSMC-IM (Wang et al., 2020). MSMC-IM uses output from MSMC (Schiffels & Durbin, 2014) to fit an isolation with migration model to coalescent rates estimated within and among two populations (Wang et al., 2020). We performed MSMC-IM analyses separately for each species. We initially tried to run MSMC with all sampled individuals, but we were unable to get our runs to complete due to computational limitations. We therefore chose two random individuals per species per population ($N = 4$ total for each species except the Abyssinian Catbird where $N = 3$). Here, we used WhatsHap (Martin et al., 2016) to phase genotypes for all individuals. We chose this method because it uses read-based phasing and does not require large sample sizes or reference SNP panels for phasing. Phasing is required for these analyses because MSMC-IM requires estimating cross-coalescence rates between at least two individuals (i.e., [?] four chromosomes). This is in contrast to MSMC demographic analyses with a single individual that do not require phasing between the two chromosomes sampled in a single diploid individual. We masked all regions with sequencing coverage lower than six in any individual to minimize inclusion of sites with phasing or genotyping errors. We ran MSMC with up to 20 iterations and 23 distinct time segments, as with individual-based demographic histories. Recent within-population histories were qualitatively similar to those estimated from single individuals but with less temporal resolution (results not shown). Less resolution may be expected since we are masking any genomic regions with low sequencing coverage in any individual, and therefore sampling less of the genome. Regardless, our main goal here was to estimate whether all our focal species shared the same patterns of population divergence and migration through time, and small shifts in overall resolution would not heavily impact these types of interpretations. For each species, we used the MSMC output as input for MSMC-IM to fit an isolation with migration model. In MSMC-IM, we used estimated empirical mutation rates for each species, and the recommended regularization settings for estimating migration rates and population sizes ($1e-8$ and $1e-6$, respectively).

RESULTS

Dataset characteristics

We sequenced 16.3 to 50.3 Gbp (mean = 33.7 Gbp) per individual, with 14.8 to 44.5 Gbp (mean = 30.1 Gbp) passing quality filtering (Table S1). Because we used a relatively divergent reference genome—not in the same genus as any of the study species—parts of the genome had no alignment coverage (range = 13.3 to 23.4% unaligned reference genome per individual) (Table S1). Overall, the data quality and quantity resulted in 6.4x to 23.9x genomic coverage per individual (mean = 14.5x) (Fig. S1), and after filtering low coverage and low-quality genomic regions, all individuals were genotyped in ~400 to 600 million sites each (Table S1).

Genomic diversity and structure

We investigated whether gene flow would be limited between populations on either side of the GRV by examining genetic variation in eastern and western populations of six bird species. The six species showed clearly different magnitudes of genetic diversity (Table 1), with estimates of observed heterozygosity ranging from 0.0008 to 0.0047 across individuals. Generally, genomic diversity was consistent within each species (Table 1; Fig. 3) with the exception of the Brown-rumped Seedeater which exhibited relatively lower diversity in individuals from the east side of the GRV.

Similar to diversity, genomic differentiation also strongly varied between species (Table 1; Fig. 5). Three species had generally low estimates of genome wide F_{ST} , with estimates below 0.06: the Abyssinian Slaty Fly-

catcher, the Abyssinian Thrush, and the Ethiopian White-eye. Two species with either very low (Abyssinian Catbirds) or inconsistent patterns (Brown-rumped Seedeater) of genomic diversity exhibited the highest values of F_{ST} . Estimates of between population diversity (also known as absolute genetic differentiation) as measured by D_{XY} also showed large variability between species (Table 1) and is largely correlated with mean observed heterozygosity with both eastern ($r = 0.71$) and western ($r = 0.85$) populations. D_{XY} is a measure of diversity proportional to the time to the most recent common ancestor (TMRCA), and is akin to nucleotide diversity estimates within a populations. Generally, estimates of F_{ST} and D_{XY} were consistent across the entire genome (Fig. S2, Fig. S3), suggesting that our mean estimates from the entire genome should be reasonably accurate. However, there were a few outlier regions for both statistics; first, F_{ST} estimates were slightly higher on the Z chromosome, and second, part of the 4A chromosome shows elevated values of D_{XY} in both the Abyssinian Catbird and the Ethiopian White-eye. These increased relative values of D_{XY} in these two taxa are expected, as part of the Zebra Finch chromosome 4A has been shown to be part of a neo-sex chromosome region in the Sylvioidea clade (Leroy et al., 2019) *sensu* Oliveros and colleagues (2019). We may expect higher genetic differentiation on sex chromosomes due to their lower relative effective population size compared to autosomes and the propensity for sex chromosomes to accumulate structural variants (e.g., Hooper, Griffith, & Price, 2018).

Regardless of substantial variation in F_{ST} , genetic structure as estimated in STRUCTURE solidly differentiated individuals from either side of the GRV (Fig. 2). Only a single individual of Ruppell’s Robin-chat showed any probability of ancestry not to its respective genetic cluster. Additionally, the phylogenomic results corroborated population differentiation between populations separated by the GRV, but with varying branch lengths to the MRCA of both populations (Fig. 2). These results suggest that the genome-wide information used here exhibits strong diagnostic power even when differentiation between populations is low.

Demographic history

Different demographic histories may lead to divergent differentiation and diversity statistics. We examined whether differences in the magnitude of drift (e.g., due to smaller population sizes) could explain differentiation across species. Demographic histories as estimated in MSMC2 largely varied by species, and in some cases populations, but was largely consistent within populations of each species (Fig. 2D). Contemporary estimated effective population sizes varied between ~50,000 to 600,000. The Abyssinian Catbird generally had a low effective population size throughout the last 200,000 years. In contrast, the Abyssinian Thrush has maintained effective populations sizes $> 400,000$ throughout the last 200,000 years. The Ruppell’s Robin-chat is the only species sampled here with large-scale fluctuations in estimated population sizes, with large increases in population sizes from ~50 to 80 kya. The Ethiopian White-eye exhibited slight changes in population size over the past 200,000 years, with estimated N_E varying between 200,000 and 400,000. Lastly, the populations of the Brown-rumped Seedeater and the Abyssinian Slaty Flycatcher exhibited distinct population histories on either side of the GRV. Generally, within individual bootstrap replicates showed slight variation in population sizes through time but generally corroborated the full dataset results (Fig. S4). Overall, there seem to be no general patterns of demographic history through time across species, with most trends appearing to be species-specific.

Population differentiation and connectivity histories

We used MSMC-IM to interpret general patterns of population differentiation and connectivity for populations separated by the GRV in each of the six focal taxa. We found that a majority of the taxa exhibited a single period of population differentiation in the past 250,000 years followed by isolation (Fig. 4). Two species showed variations of this pattern: the Abyssinian Catbird and the Abyssinian Thrush. The isolation with migration model for the Abyssinian Catbird indicated a prolonged period of population differentiation with ongoing connectivity between populations (Fig. 4). For the Abyssinian Thrush, the model implied an initial period of differentiation followed by several bouts of connectivity (i.e., gene flow) through time (Fig. 4). This pattern in the Abyssinian Thrush is consistent with very low F_{ST} estimates (Table 1), a very shallow MRCA in the phylogenomic results (Fig. 2), and shared mtDNA haplotypes among populations on

either side of the GRV (Fig. 2).

Because of caveats associated with a lack of knowledge about and estimating generation times in tropical montane birds (see next section of RESULTS), we refrain from interpreting exact divergence times, but rather interpret the general trends of population differentiation and connectivity through time. For example, four of the species (Ruppell’s Robin-chat, Brown-rumped Seedeater, Abyssinian Slaty Flycatcher, Ethiopian White-eye) exhibited a single period of population differentiation followed by isolation (Fig. 4). If we strictly interpret the time periods of these differentiation events with assumed generation times, we would interpret that these species generally exhibit asynchronous diversification. However, with any error in our generation time assumptions, these periods of differentiation could potentially be synchronous. Because of these issues, we only discuss the general trends of population differentiation and connectivity through time.

Demographic history and isolation with migration model caveats

Estimating population sizes and timing demographic events depends on accurate estimates of mutation rate per generation. As such, an important caveat to the demographic and isolation with migration results relates to assumptions about generation times and mutation rates in these bird species. Generation times are reliable when information about age of sexual maturity, adult survival, and offspring survival is known in a species. In tropical bird species, little of this information is known, and indeed none of the focal species studied here have known generation times. Additionally, reproductive effort is known to vary among temperate and tropical species (Hau et al., 2010), and even along elevational gradients in tropical regions (Boyce et al., 2015).

The mutation rates we used here were calculated from time-calibrated avian phylogenies, and implicitly rely on accurate timing of nodes from the estimated phylogeny (Oliveros et al., 2019). These mutation rate estimates are in numbers of mutations per year (Table 1) and are generally consistent in magnitude with other Passeriformes mutation rate estimates (Nam et al., 2010). As such, in the context of estimating mutation rate per generation, we likely have more certainty in our absolute mutation rates (per year) and less certainty in mutation rate per generation because of unknown generation times. Regardless, we expect that our two-year generation time assumptions are within a reasonable amount of error from real values, e.g., we don’t expect these values to be off by an order of magnitude.

Overall, given the uncertainties with both mutation rate and generation times, we should expect that any population sizes and timings estimated in demographic histories and isolation with migration models may vary. Absolute timing events will be inaccurate proportional to error in our absolute mutation rate estimates (mutations per year). Because our population size estimates (i.e., $4N\mu$) directly depend on mutation rates per generation, the error in population size estimates will directly scale with our error in estimating how the mutation rate relates to generation times. For example, if the true generation times of our focal species range from a year to four years (and we used an estimate of two years), estimated effective population sizes will be doubled or halved, respectively. Because we see a strong correlation between genomic diversity and harmonic mean of population sizes including all individuals across species (Fig. 3), even with potential error in estimating generation times, we can still interpret relative differences between species’ genomic diversities and population size trends.

Correlates of genetic diversity and differentiation with species’ characteristics

Genomic diversity across individuals has a strong positive relationship with the harmonic mean of effective population size estimated from each individual (Fig. 3A; $R^2 = 0.689$). This association remained when we took evolutionary history into account with phylogenetic independent contrasts (Fig. 3C; $R^2 = 0.599$). We may expect these patterns *a priori* because genomic diversity and population sizes should be directly linked. In contrast, we found no relationship between genomic diversity and HWI measurements across individuals (Fig. 3B, Fig. 3D).

Genomic differentiation (as measured by F_{ST}) showed a significant negative association with dispersal ability (measured by HWI) (Fig. 5A). This relationship remained when accounting for evolutionary history with

PICs (Fig. 5C). Genomic differentiation showed a negative non-significant negative association with harmonic mean of effective population size (Fig. 5B, Fig. 5D).

DISCUSSION

The Great Rift Valley is a biogeographic barrier to montane birds.

In all six species examined here, we were able to differentiate populations on either side of the GRV using both genetic clustering and phylogenetic methods (Fig. 2). Based on these results we may infer that the GRV has at least acted as a moderate barrier to gene flow in montane birds, even if they have a generally higher dispersal ability than non-flying species. We found that genetic differentiation (as measured by F_{ST}) varies by an order of magnitude across species (Table 1), and isolation with migration models with varied connectivity patterns over the past ~350,000 years indicate population differentiation across the GRV was at least partly species-specific. But which species-specific factors may have driven these differences in the magnitude of population differentiation?

Dispersal ability shows a significant negative association with a measure of population genetic differentiation (Fig. 5A, Fig. 5C). It is interesting to find this significant relationship even with a small number of focal taxa. It is intuitive that dispersal ability would in general have a negative association with the ability for fragmented populations to differentiate, as increased dispersal among populations would generally lead to increased gene flow between populations. As such, different modes of dispersal show different tempos of population differentiation (Medina, Cooke, & Ord, 2018), and dispersal ability has been implicated in (1) differential rates of population differentiation in closely-related taxa (e.g., Peterman et al., 2015), and (2) differential macroevolutionary trends of speciation rates (e.g., Claramunt et al., 2011; Weeks & Claramunt, 2014). Our results in Ethiopian montane forest birds are suggestive of a relationship between dispersal ability and geographic population structure and is an intriguing line of inquiry for future study.

Population sizes show a weak negative association with population genetic differentiation (not statistically significant; Fig. 5B, Fig. 5D). This is somewhat expected, as smaller and more fragmented populations will naturally exhibit less gene flow and undergo relatively faster genetic drift. Additionally, the calculations for genetic differentiation (as measured by F_{ST}) intrinsically incorporate within-population genetic diversity, which is generally linked with long-term effective population sizes (Fig. 3A).

Comparative patterns of population structure.

Montane Ethiopian species share at least four biogeographic barriers (briefly reviewed by Manthey et al., 2017), with the GRV being the most studied to date. A diversity of taxa exhibit population or species level differentiation on either side of the GRV, including brush-furred rats, groove-toothed rats, Ethiopian wolf, gelada, frogs of several genera, giant lobelia, coffee, and the species studied here (Belay & Mori, 2006; Evans et al., 2011; Freilich et al., 2016; Gottelli et al., 2004; Kebede et al., 2007; Komarova et al., 2021; Manthey et al., 2017; Reyes-Velasco et al., 2018; Reyes-Velasco et al., 2018; Silvestrini et al., 2007).

While the GRV has impacted a plethora of taxa, species' evolutionary responses have varied in timing and/or magnitude. For example, the Ethiopian wolf (*Canis simensis*) is a relatively good disperser and exhibits weakly differentiated populations on either side of the GRV; this differentiation is hypothesized to have occurred over the last 20,000 years since the Last Glacial Maximum of the Pleistocene Epoch (Gottelli et al., 2004). In contrast, semi-arboreal frogs (Genus *Leptopelis*) are relatively poor dispersers and exhibit many distinct lineages across the Ethiopian Highlands with divergence dates predating the Pleistocene (Reyes-Velasco et al., 2018). Based on this study, songbirds have diversified during Pleistocene glacial cycles, with population differentiation and cessation of connectivity dating to the past 50 kya to 350 kya (Fig. 4). Interestingly, population differentiation in most taxa appears to have occurred during a pulse Great Rift Valley volcanism that took place between 170 kya and 320 kya (Hutchison et al., 2016).

The asynchronous and varied patterns of diversification in Ethiopian montane taxa exemplify the complex geological and climatic history of the region. The formation of the GRV is old (~20 Mya) and ongoing

(Frisch et al., 2010; Sepulchre et al., 2006), which has provided growing and continuous elevational heterogeneity in the central portion of the Ethiopian Highlands. During the last few million years, the Pleistocene climatic cycles have added climatic heterogeneity to the topographic variation; during glacial minima (e.g., contemporary periods), the GRV was likely relatively arid and hot compared to glacial maxima (e.g., the Last Glacial Maximum). These oscillations would have generally resulted in elevational shifts and greater potential geographic range in montane taxa during glacial maxima. Montane species inhabiting the Ethiopian Highlands would have experienced these combined geological and climatic forces; their population differentiation throughout these periods would have likely been shaped by the age of the lineage inhabiting the highlands (i.e., time of colonization) and species-specific dispersal ability, niche breadth, and interpopulation reproductive compatibility shaping whether populations underwent gene flow during periods of relatively higher habitat connectivity.

Drivers of genomic diversity. Genomic diversity is expected to be influenced by mutation rates, patterns of selection and linked selection, and effective population sizes (Amos & Harwood, 1998; Ellegren & Galtier, 2016). The largest effectors of mutation rates and effective population sizes are life history characteristics, namely fecundity rates (Romiguier et al., 2014). When focusing on species with similar life history strategies—as we do here with passerine birds—we may expect most of the variation in genomic diversity to be shaped by patterns of long-term linked selection across the genome and the demographic histories of populations that in turn shape effective population sizes. We find a negative correlation between genome wide estimates of F_{ST} and D_{XY} , indicative of the effects of widespread linked selection in all six taxa (Fig. S5) (Cruickshank & Hahn, 2014), although with our limited sample sizes it wouldn't be practical to identify direct effects of linked selection on genomic diversity. Overall, we find that the strongest predictors of genomic diversity across Ethiopian montane bird populations are recent population sizes and harmonic mean of population sizes over the past 200,000 years (Fig. 3).

CONCLUSIONS

We sequenced whole genomes at low to moderate coverage for two populations each of six montane forest bird species, to test if the Great Rift Valley (GRV) is a barrier to dispersal in a taxonomic group with relatively high dispersal ability. We found that we were able to distinguish individuals from either side of the GRV using genetic clustering and phylogenomic methods. The magnitude of genetic differentiation varied by species, with relatively poorer dispersers generally exhibiting higher amounts of genetic differentiation. Additionally, isolation with migration models indicated varied patterns of connectivity and divergence over the past 350,000 years among these six taxa. We found that each species generally had unique demographic histories, with some species exhibiting different demographic histories on either side of the GRV. Lastly, we found that demographic history explained more than 80% of the variance in genomic diversity across samples, while dispersal ability was not correlated with genomic diversity.

ACKNOWLEDGEMENTS

We thank the Ethiopian Wildlife Conservation Authority and the Oromia Forest and Wildlife Enterprise for assistance and allowing us to conduct field work. We thank Patrick Bradley and Luke Campillo for fieldwork help. We thank Rock Hewn Tours personnel, especially Megersa Kelbessa, for transportation, translation, and field assistance. We thank the National and State Park Rangers and Guides with support while in the field. Yvonne and Guy Levene at Bale Mountain Lodge provided assistance while in Bale Mountain National Park. This work was supported by the National Science Foundation under Grant #1953688 to JDM, Texas Tech University startup funds to JDM, and New York University Abu Dhabi (NYUAD) research funds AD180 to SB.

AUTHOR CONTRIBUTIONS

All authors contributed to the project ideas and manuscript completion and editing. JDM and SB contributed funding. JDM & YB performed fieldwork. JDM performed DNA extractions, bioinformatics, and wrote the first draft of the manuscript.

DATA ACCESSIBILITY

All raw sequencing data is uploaded to NCBI’s Sequence Read Archive (SRA) accessioned under BioProject PRJNA605410. All code used for data analysis in this project can be found at: github.com/jdmanthey/ethiopia_grv_birds. Dataset citation:

[dataset] Manthey JD, Bourgeois Y, Yonas M, Boissinot S. Ethiopia Highland Birds GRV whole genome sequencing. National Center for Biotechnology Information Sequence Read Archive (SRA). <https://www.ncbi.nlm.nih.gov/bioproject/PRJNA605410/>

Figure 1. Sampling geographic context in Ethiopia. We sampled three sites in the Menagesha Forest Preserve west of the Great Rift Valley (GRV) and two sites in the Bale Mountains National Park east of the GRV: Katcha and Dinsho. See Table S1 for geographic coordinates of each site. The Blue Nile Valley is also labeled here for reference because it is mentioned multiple times in the text, although we are not directly testing anything with that barrier in this study. Note that the Ethiopian Highlands are often defined as starting at different elevations (e.g., 2000 m). Here, we shaded regions greater than 1800 m as some regions have montane forests lower than 2000 m. All photos by JDM.

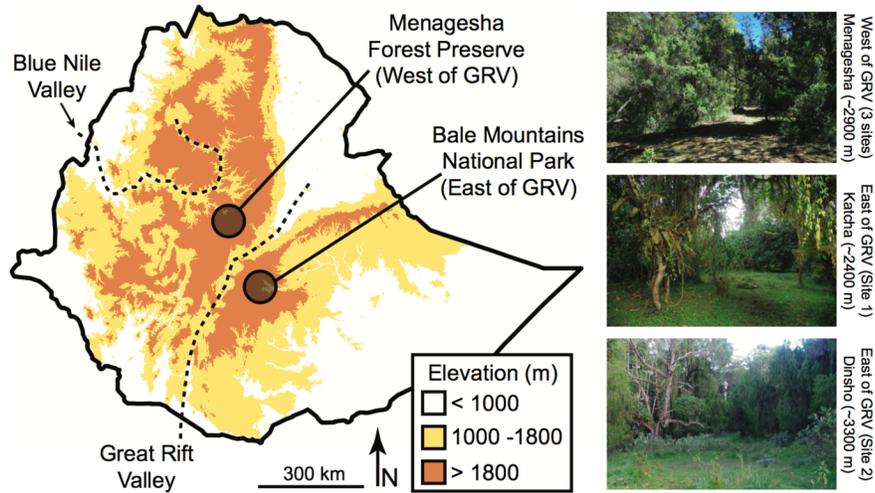


Figure 2. Evolutionary and demographic history in six Ethiopian montane birds. (A) Inference of genetic structure using the program STRUCTURE for each species. Each column indicates one individual and the proportion of each color the probability of assignment to one of two genetic clusters. Individuals sampled to the west or east of the Great Rift Valley (GRV) are indicated on the left and right of the dotted line, respectively. (B) Species tree topologies for each species estimated in ASTRAL. (C) ND2 mitochondrial DNA haplotype networks. (D) Estimates of demographic history for each individual estimated with MSMC. Complete results including bootstraps are presented in Fig. S4. Different individuals are represented with differently colored lines for increased clarity. Plots are separated for individuals west (left) and east (right) of the GRV. General elevational affinities shown for each species in the bottom corner of photos. All photos by JDM.

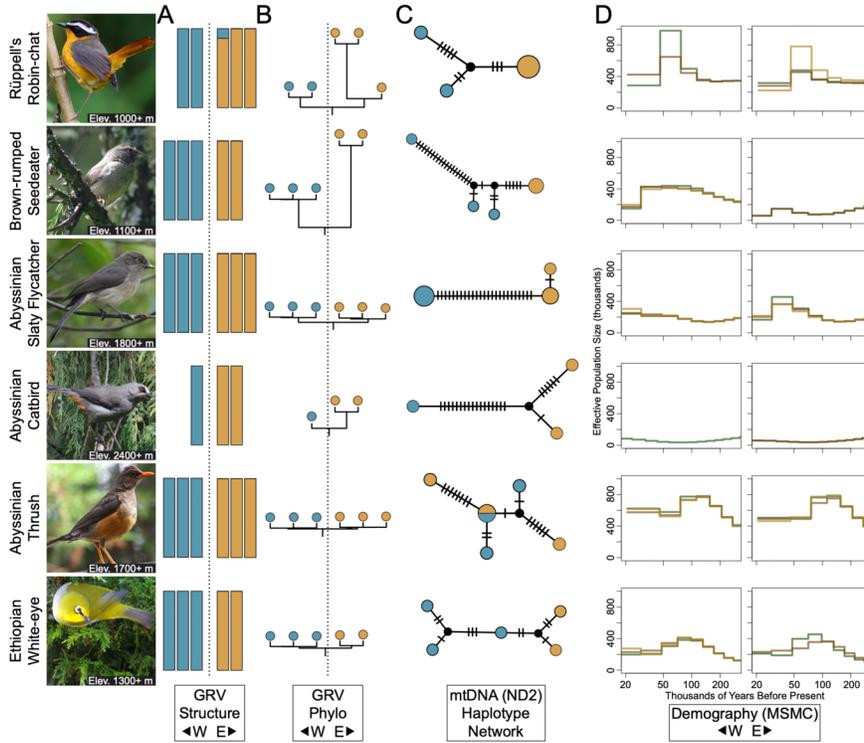


Figure 3. Relationships of genomic diversity and species-specific characteristics. (A + C) Relationship of genomic diversity and harmonic mean population sizes through time estimated in MSMC showing (A) raw data and (C) phylogenetic independent contrasts (PIC). (B + D) Relationship of genomic diversity and hand-wing index, a proxy measurement for dispersal ability showing (B) raw data and (D) PIC.

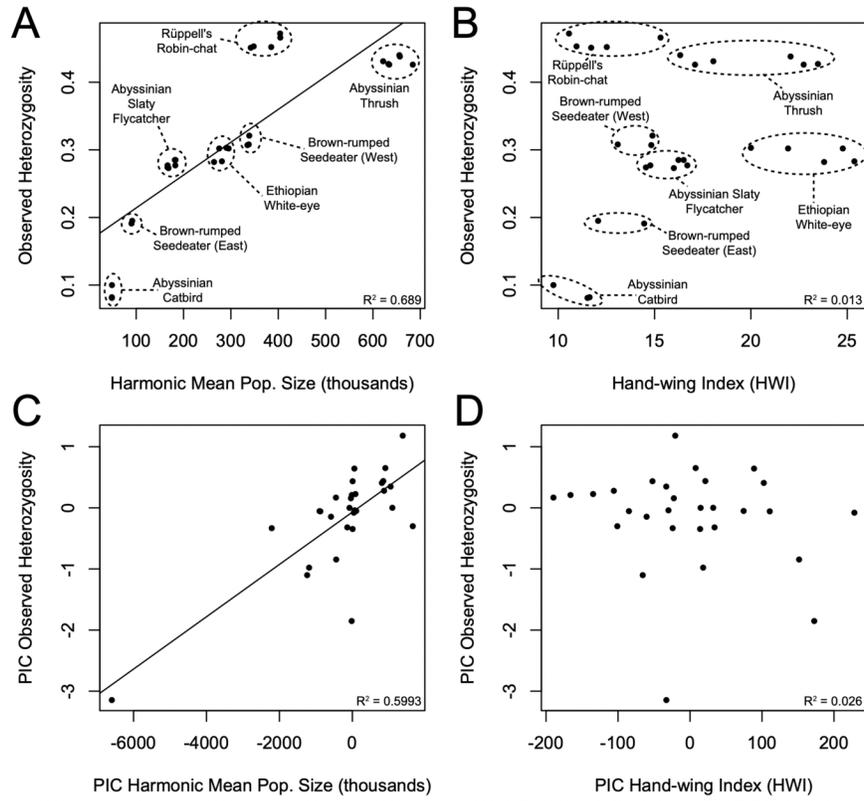


Figure 4. Isolation with migration models as estimated in MSMC-IM. On left, estimates of migration rates through time for each species. On right, the cumulative migration probabilities throughout the past 500,000 years. In species with a single peak on the left and a steep slope on the right, we would infer a single period of differentiation followed by isolation. In contrast, species with multiple or extended peaks on left plots indicate multiple or prolonged bouts of gene flow between differentiating populations. Population differentiation is inferred to be occurring when $0 < M(t) < 1$.

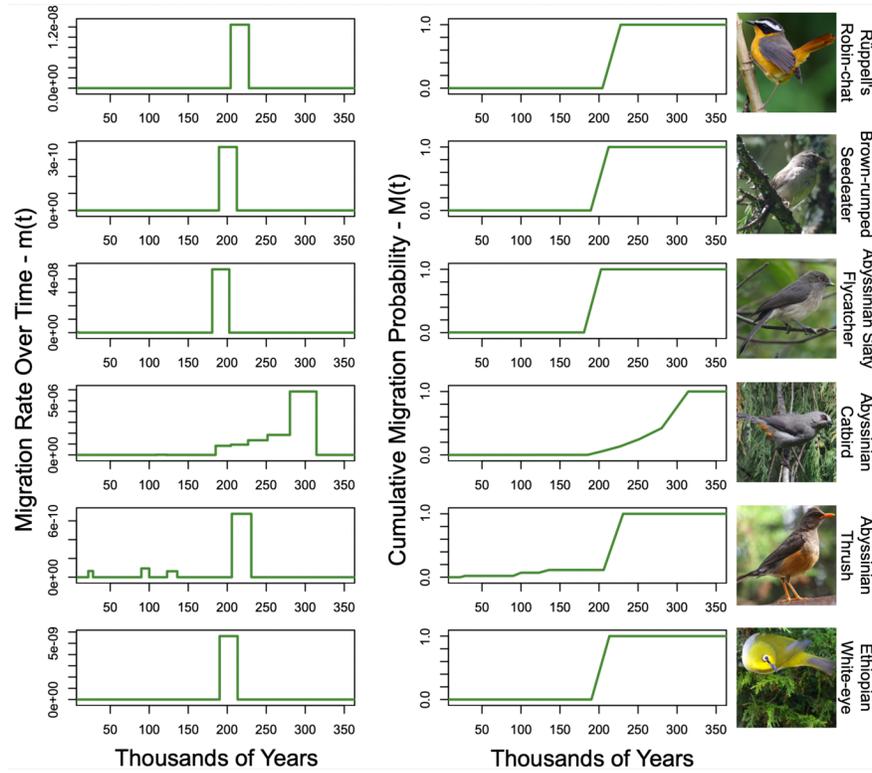


Figure 5. Relationship of genomic differentiation as measured by F_{ST} and species-specific characteristics. (A + C) Relationship of F_{ST} and hand-wing index showing (A) raw data and (C) phylogenetic independent contrasts (PIC). (B + D) Relationship of F_{ST} and harmonic mean population sizes through time showing (B) raw data and (D) PIC. Genomic differentiation as measured by F_{ST} has no strong linear relationship with measured characteristics of each species. Hand-wing index and population sizes were included as the mean value for all individuals per species that were sampled for this study.

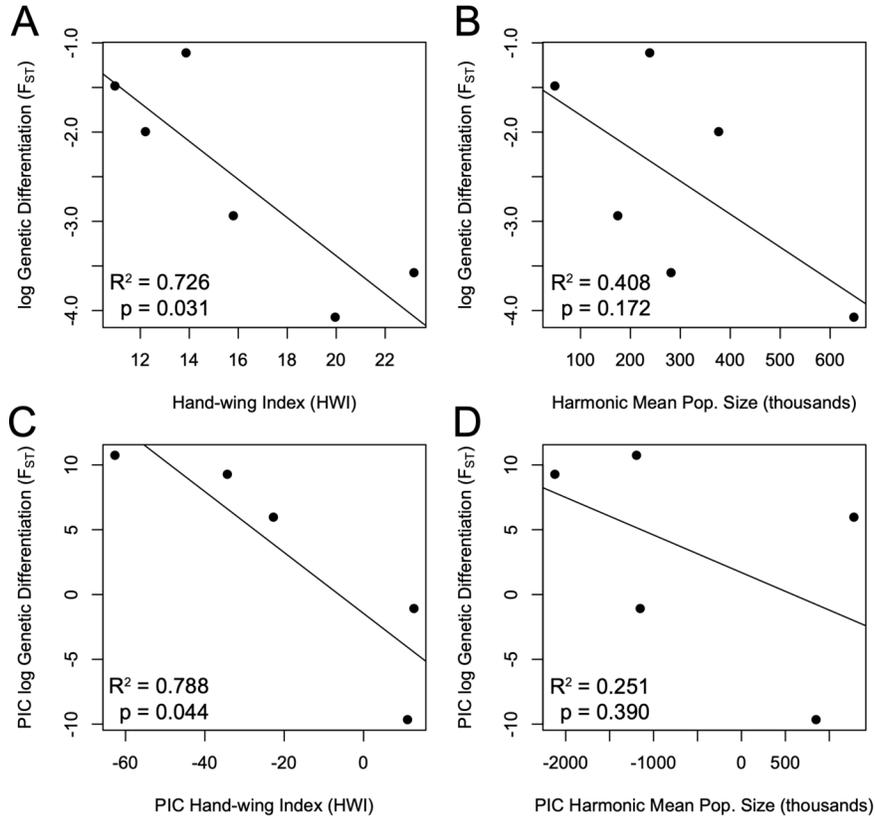


Table 1. Summary statistics for each species included in this study.

Species	Common name	N West	N East	Mean HWI	F _{ST}	D _{XY} ¹	# SNPs ²	# total sites ²	H _O West ¹	H _O East ¹
<i>Cossypha semirufa</i>	Rüppell's Robin-chat	2	3	12.21	0.136	0.443	6.24	509	0.469	0.453
<i>Crithagra tris-triata</i>	Brown-rumped Seedeater	3	2	13.87	0.329	0.321	4.61	557	0.312	0.193
<i>Melaenornis choco-latinus</i>	Abyssinian Slaty Flycatcher	3	3	15.80	0.053	0.239	3.95	518	0.275	0.283
<i>Sylvia galinieri</i>	Abyssinian Catbird	1	2	10.96	0.227	0.079	0.61	354	0.082	0.091
<i>Turdus abyssinicus</i>	Abyssinian Thrush	3	3	19.96	0.017	0.247	1.75	203	0.435	0.428
<i>Zosterops polio-gastrus</i>	Ethiopian White-eye	3	2	23.17	0.028	0.239	3.44	443	0.303	0.283

N = sam- ple size, HWI = hand- wing index, H _O = ob- served het- erozy- gosity, F _{ST} and D _{XY} = mea- sures of ge- netic differ- entia- tion, # SNPs = num- ber of SNPs geno- typed per species, # total sites = total num- ber of sites geno- typed in every indi- vidual (i.e., no miss- ing data) and pass- ing filters for	N = sam- ple size, HWI = hand- wing index, H _O = ob- served het- erozy- gosity, F _{ST} and D _{XY} = mea- sures of ge- netic differ- entia- tion, # SNPs = num- ber of SNPs geno- typed per species, # total sites = total num- ber of sites geno- typed in every indi- vidual (i.e., no miss- ing data) and pass- ing filters for	N = sam- ple size, HWI = hand- wing index, H _O = ob- served het- erozy- gosity, F _{ST} and D _{XY} = mea- sures of ge- netic differ- entia- tion, # SNPs = num- ber of SNPs geno- typed per species, # total sites = total num- ber of sites geno- typed in every indi- vidual (i.e., no miss- ing data) and pass- ing filters for	N = sam- ple size, HWI = hand- wing index, H _O = ob- served het- erozy- gosity, F _{ST} and D _{XY} = mea- sures of ge- netic differ- entia- tion, # SNPs = num- ber of SNPs geno- typed per species, # total sites = total num- ber of sites geno- typed in every indi- vidual (i.e., no miss- ing data) and pass- ing filters for	N = sam- ple size, HWI = hand- wing index, H _O = ob- served het- erozy- gosity, F _{ST} and D _{XY} = mea- sures of ge- netic differ- entia- tion, # SNPs = num- ber of SNPs geno- typed per species, # total sites = total num- ber of sites geno- typed in every indi- vidual (i.e., no miss- ing data) and pass- ing filters for	N = sam- ple size, HWI = hand- wing index, H _O = ob- served het- erozy- gosity, F _{ST} and D _{XY} = mea- sures of ge- netic differ- entia- tion, # SNPs = num- ber of SNPs geno- typed per species, # total sites = total num- ber of sites geno- typed in every indi- vidual (i.e., no miss- ing data) and pass- ing filters for	N = sam- ple size, HWI = hand- wing index, H _O = ob- served het- erozy- gosity, F _{ST} and D _{XY} = mea- sures of ge- netic differ- entia- tion, # SNPs = num- ber of SNPs geno- typed per species, # total sites = total num- ber of sites geno- typed in every indi- vidual (i.e., no miss- ing data) and pass- ing filters for	N = sam- ple size, HWI = hand- wing index, H _O = ob- served het- erozy- gosity, F _{ST} and D _{XY} = mea- sures of ge- netic differ- entia- tion, # SNPs = num- ber of SNPs geno- typed per species, # total sites = total num- ber of sites geno- typed in every indi- vidual (i.e., no miss- ing data) and pass- ing filters for	N = sam- ple size, HWI = hand- wing index, H _O = ob- served het- erozy- gosity, F _{ST} and D _{XY} = mea- sures of ge- netic differ- entia- tion, # SNPs = num- ber of SNPs geno- typed per species, # total sites = total num- ber of sites geno- typed in every indi- vidual (i.e., no miss- ing data) and pass- ing filters for	N = sam- ple size, HWI = hand- wing index, H _O = ob- served het- erozy- gosity, F _{ST} and D _{XY} = mea- sures of ge- netic differ- entia- tion, # SNPs = num- ber of SNPs geno- typed per species, # total sites = total num- ber of sites geno- typed in every indi- vidual (i.e., no miss- ing data) and pass- ing filters for	N = sam- ple size, HWI = hand- wing index, H _O = ob- served het- erozy- gosity, F _{ST} and D _{XY} = mea- sures of ge- netic differ- entia- tion, # SNPs = num- ber of SNPs geno- typed per species, # total sites = total num- ber of sites geno- typed in every indi- vidual (i.e., no miss- ing data) and pass- ing filters for
---	---	---	---	---	---	---	---	---	---	---

LITERATURE CITED

- [dataset] Manthey JD, Bourgeois Y, Yonas M, Boissinot S. Ethiopia Highland Birds GRV whole genome sequencing. National Center for Biotechnology Information Sequence Read Archive (SRA). <https://www.ncbi.nlm.nih.gov/bioproject/PRJNA605410/>
- Amos, W., & Harwood, J. (1998). Factors affecting levels of genetic diversity in natural populations. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 353 (1366), 177-186.
- Bandelt, H.-J., Forster, P., & Röhl, A. (1999). Median-joining networks for inferring intraspecific phylogenies. *Molecular Biology and Evolution*, 16 (1), 37-48.
- Barber, B. R., & Klicka, J. (2010). Two pulses of diversification across the Isthmus of Tehuantepec in a montane Mexican bird fauna. *Proceedings of the Royal Society of London, Series B: Biological Sciences*, rspb20100343.
- Belay, G., & Mori, A. (2006). Intraspecific phylogeographic mitochondrial DNA (D-loop) variation of Gelada baboon, *Theropithecus gelada*, in Ethiopia. *Biochemical Systematics and Ecology*, 34 (7), 554-561.
- Boyce, A. J., Freeman, B. G., Mitchell, A. E., & Martin, T. E. (2015). Clutch size declines with elevation in tropical birds. *The Auk: Ornithological Advances*, 132 (2), 424-432.
- Bryja, J., Kostin, D., Meheretu, Y., Šumbera, R., Bryjová, A., Kasso, M., . . . Lavrenchenko, L. A. (2018). Reticulate Pleistocene evolution of Ethiopian rodent genus along remarkable altitudinal gradient. *Molecular Phylogenetics and Evolution*, 118, 75-87.
- Bushnell, B. (2014). *BBMap: A Fast, Accurate, Splice-Aware Aligner*. Retrieved from
- Charif, D., & Lobry, J. R. (2007). SeqinR 1.0-2: a contributed package to the R project for statistical computing devoted to biological sequences retrieval and analysis. In *Structural Approaches to Sequence Evolution* (pp. 207-232): Springer.
- Chikhi, L., Rodríguez, W., Grusea, S., Santos, P., Boitard, S., & Mazet, O. (2018). The IICR (inverse instantaneous coalescence rate) as a summary of genomic diversity: insights into demographic inference and model choice. *Heredity*, 120 (1), 13-24.
- Claramunt, S., Derryberry, E. P., Remsen Jr, J., & Brumfield, R. T. (2011). High dispersal ability inhibits speciation in a continental radiation of passerine birds. *Proceedings of the Royal Society B: Biological Sciences*, 279 (1733), 1567-1574.
- Clement, P. (2020). Brown-rumped Seedeater (*Crithagra tristriata*), version 1.0. In Birds of the World (J. del Hoyo, A. Elliott, J. Sargatal, D. A. Christie, and E. de Juana, Editors). Cornell Lab of Ornithology, Ithaca, NY, USA. <https://doi.org/10.2173/bow.brrsee1.01>.
- Collar, N. (2020). Rüppell's Robin-Chat (*Cossypha semirufa*), version 1.0. In Birds of the World (J. del Hoyo, A. Elliott, J. Sargatal, D. A. Christie, and E. de Juana, Editors). Cornell Lab of Ornithology, Ithaca, NY, USA. <https://doi.org/10.2173/bow.rurcha1.01>.
- Collar, N., & Robson, C. (2021). Abyssinian Catbird (*Sylvia galinieri*), version 1.1. In Birds of the World (J. del Hoyo, A. Elliott, J. Sargatal, D. A. Christie, and E. de Juana, Editors). Cornell Lab of Ornithology, Ithaca, NY, USA. <https://doi.org/10.2173/bow.abycat1.01.1>.
- Cruickshank, T. E., & Hahn, M. W. (2014). Reanalysis suggests that genomic islands of speciation are due to reduced diversity, not reduced gene flow. *Molecular Ecology*, 23 (13), 3133-3157.
- Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., . . . Sherry, S. T. (2011). The variant call format and VCFtools. *Bioinformatics*, 27 (15), 2156-2158.
- Darriba, D., Taboada, G. L., Doallo, R., & Posada, D. (2012). jModelTest 2: more models, new heuristics and parallel computing. *Nature Methods*, 9 (8), 772.

- del Hoyo, J., Collar, N., & Kirwan, G. (2020). Abyssinian Thrush (*Turdus abyssinicus*), version 1.0. In *Birds of the World* (J. del Hoyo, A. Elliott, J. Sargatal, D. A. Christie, and E. de Juana, Editors). Cornell Lab of Ornithology, Ithaca, NY, USA. <https://doi.org/10.2173/bow.abythr1.01>.
- Derjushева, S., Kurganova, A., Habermann, F., & Gaginskaya, E. (2004). High chromosome conservation detected by comparative chromosome painting in chicken, pigeon and passerine birds. *Chromosome Research*, *12* (7), 715.
- Dessie, G., & Kleman, J. (2007). Pattern and magnitude of deforestation in the South Central Rift Valley Region of Ethiopia. *Mountain research and development*, *27* (2), 162-169.
- Ellegren, H., & Galtier, N. (2016). Determinants of genetic diversity. *Nature reviews genetics*, *17* (7), 422.
- Evans, B. J., Bliss, S. M., Mendel, S. A., & Tinsley, R. C. (2011). The Rift Valley is a major barrier to dispersal of African clawed frogs (*Xenopus*) in Ethiopia. *Molecular Ecology*, *20* (20), 4216-4230.
- Felsenstein, J. (1985). Phylogenies and the comparative method. *The American Naturalist*, *125* (1), 1-15.
- Freilich, X., Anadón, J. D., Bukala, J., Calderon, O., Chakraborty, R., & Boissinot, S. (2016). Comparative Phylogeography of Ethiopian anurans: impact of the Great Rift Valley and Pleistocene climate change. *BMC Evolutionary Biology*, *16* (1), 206.
- Friis, I., Edwards, S., Ensermu, K., & Sebsebe, D. (2001). Diversity and endemism in the flora of Ethiopia and Eritrea—what do the published flora volumes tell us. *Biol. Skr*, *54*, 173-193.
- Frisch, W., Meschede, M., & Blakey, R. C. (2010). *Plate tectonics: Continental Drift and mountain building*: Springer Science & Business Media.
- Gottelli, D., Marino, J., SILLERO-ZUBIRI, C., & Funk, S. M. (2004). The effect of the last glacial age on speciation and population genetic structure of the endangered Ethiopian wolf (*Canis simensis*). *Molecular Ecology*, *13* (8), 2275-2286.
- Griffin, D. K., Robertson, L. B., Tempest, H. G., Vignal, A., Fillon, V., Crooijmans, R. P., . . . Carre, W. (2008). Whole genome comparative studies between chicken and turkey and their implications for avian genome evolution. *BMC Genomics*, *9* (1), 168.
- Guindon, S., Dufayard, J.-F., Lefort, V., Anisimova, M., Hordijk, W., & Gascuel, O. (2010). New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Systematic Biology*, *59* (3), 307-321.
- Habel, J. C., Borghesio, L., Newmark, W. D., Day, J. J., Lens, L., Husemann, M., & Ulrich, W. (2015). Evolution along the Great Rift Valley: phenotypic and genetic differentiation of East African white-eyes (Aves, Zosteropidae). *Ecology and Evolution*, *5* (21), 4849-4862.
- Hamilton, A. J., Basset, Y., Benke, K. K., Grimbacher, P. S., Miller, S. E., Novotny, V., . . . Yen, J. D. (2010). Quantifying uncertainty in estimation of tropical arthropod species richness. *The American Naturalist*, *176* (1), 90-95.
- Hau, M., Ricklefs, R. E., Wikelski, M., Lee, K. A., & Brawn, J. D. (2010). Corticosterone, testosterone and life-history strategies of birds. *Proceedings of the Royal Society B: Biological Sciences*, *277* (1697), 3203-3212.
- Hooper, D. M., Griffith, S. C., & Price, T. D. (2018). Sex chromosome inversions enforce reproductive isolation across an avian hybrid zone. *Molecular Ecology*. doi:10.1111/mec.14874
- Hubisz, M. J., Pollard, K. S., & Siepel, A. (2010). PHAST and RPHAST: phylogenetic analysis with space/time models. *Briefings in bioinformatics*, *12* (1), 41-51.

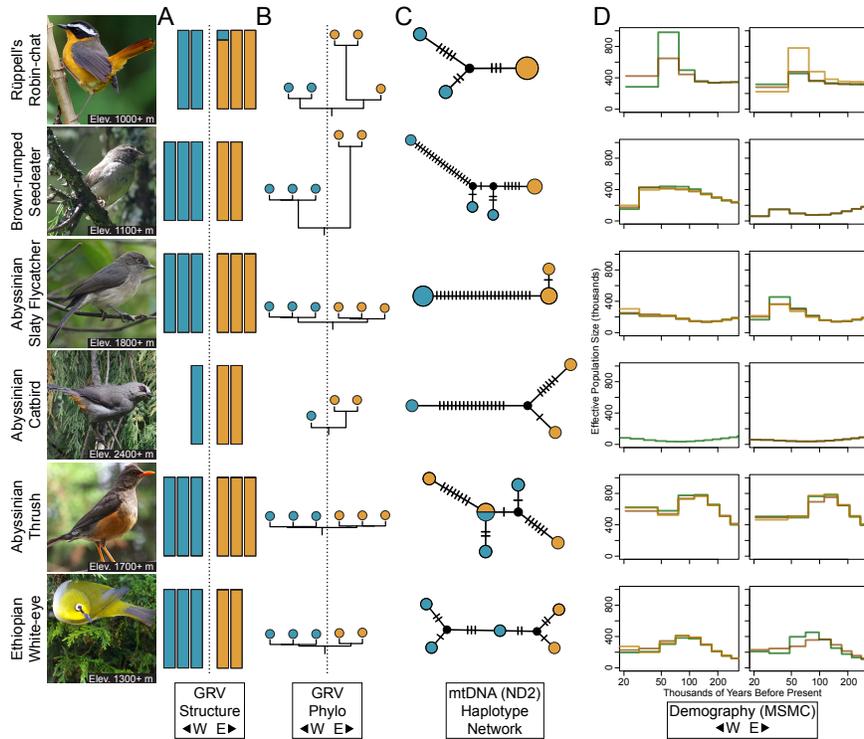
- Hutchison, W., Fusillo, R., Pyle, D. M., Mather, T. A., Blundy, J. D., Biggs, J., . . . Barfod, D. N. (2016). A pulse of mid-Pleistocene rift volcanism in Ethiopia at the dawn of modern humans. *Nature Communications*, *7* (1), 1-12.
- Kebede, M., Ehrlich, D., Taberlet, P., Nemomissa, S., & Brochmann, C. (2007). Phylogeography and conservation genetics of a giant lobelia (*Lobelia giberroa*) in Ethiopian and Tropical East African mountains. *Molecular Ecology*, *16* (6), 1233-1243.
- Kipp, F. A. (1959). Der Handflugel-Index als flugbiologisches Mass. *Die Vogelwarte*, *20* (2), 77-86.
- Komarova, V. A., Kostin, D. S., Bryja, J., Mikula, O., Bryjova, A., Čížková, D., . . . Lavrenchenko, L. A. (2021). Complex reticulate evolution of speckled brush-furred rats (*Lophuromys*) in the Ethiopian centre of endemism. *Molecular Ecology*, *30* (10), 2349-2365.
- Largen, M., & Spawls, S. (2010). *The amphibians and reptiles of Ethiopia and Eritrea* : Edition Chimaira.
- Leigh, J. W., & Bryant, D. (2015). popart: full-feature software for haplotype network construction. *Methods in Ecology and Evolution*, *6* (9), 1110-1116.
- Leroy, T., Anselmetti, A., Tilak, M. K., Berard, S., Csukonyi, L., Gabrielli, M., . . . Nabholz, B. (2019). A bird's white-eye view on neo-sex chromosome evolution. *bioRxiv peer-reviewed and recommended by PCI Evolutionary Biology*, *505610*, ver. 4 . doi:10.1101/505610
- Li, H., & Durbin, R. (2009). Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics*, *25* (14), 1754-1760.
- Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., . . . Durbin, R. (2009). The sequence alignment/map format and SAMtools. *Bioinformatics*, *25* (16), 2078-2079.
- Manthey, J. D., Reyes-Velasco, J., Freilich, X., & Boissinot, S. (2017). Diversification in a biodiversity hotspot: genomic variation in the river frog *Amietia nutti* across the Ethiopian Highlands. *Biological Journal of the Linnean Society*, *122* (4), 801-813.
- Martin, M., Patterson, M., Garg, S., Fischer, S., Pisanti, N., Klau, G. W., . . . Marschall, T. (2016). WhatsHap: fast and accurate read-based phasing. *BioRxiv DOI: 10.1101/085050* , 085050.
- Mazet, O., Rodriguez, W., Grusea, S., Boitard, S., & Chikhi, L. (2016). On the importance of being structured: instantaneous coalescence rates and human evolution—lessons for ancestral population size inference? *Heredity*, *116* (4), 362-371.
- McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., . . . Daly, M. (2010). The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Research*, *20* (9), 1297-1303.
- Medina, I., Cooke, G. M., & Ord, T. J. (2018). Walk, swim or fly? Locomotor mode predicts genetic differentiation in vertebrates. *Ecology Letters*, *21* (5), 638-645.
- Miraldo, A., Li, S., Borregaard, M. K., Florez-Rodriguez, A., Gopalakrishnan, S., Rizvanovic, M., . . . Nogues-Bravo, D. (2016). An Anthropocene map of genetic diversity. *Science*, *353* (6307), 1532-1535.
- Mora, C., Tittensor, D. P., Adl, S., Simpson, A. G., & Worm, B. (2011). How many species are there on Earth and in the ocean? *PLoS Biology*, *9* (8), e1001127.
- Musher, L. J., Ferreira, M., Auerbach, A. L., McKay, J., & Cracraft, J. (2019). Why is Amazonia a 'source' of biodiversity? Climate-mediated dispersal and synchronous speciation across the Andes in an avian group (Tityrinae). *Proceedings of the Royal Society B*, *286* (1900), 20182343.
- Myers, N., Mittermeier, R. A., Mittermeier, C. G., Da Fonseca, G. A., & Kent, J. (2000). Biodiversity hotspots for conservation priorities. *Nature*, *403* (6772), 853-858.

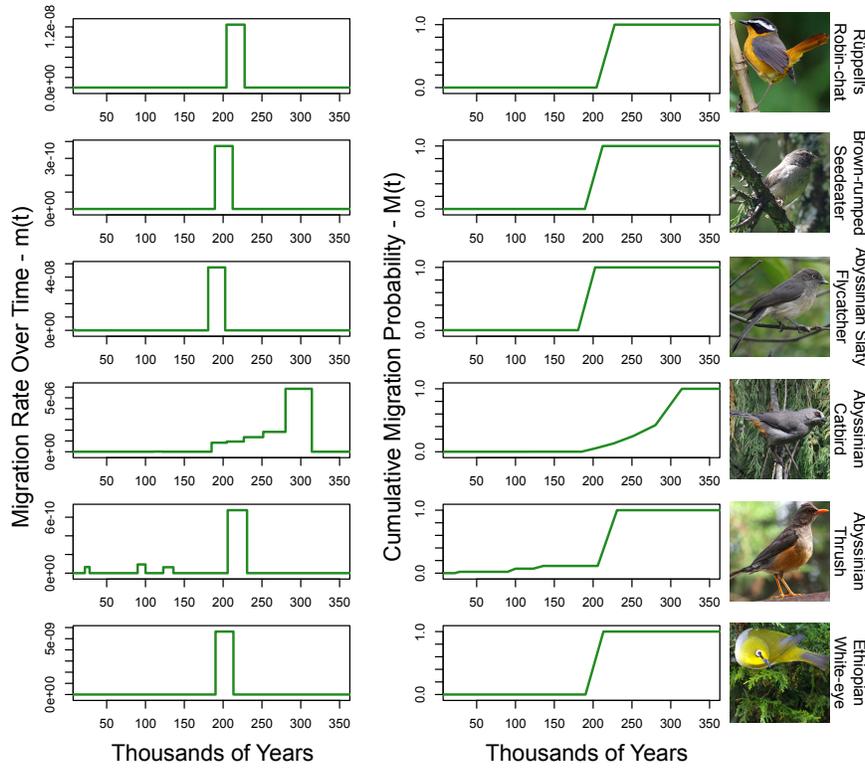
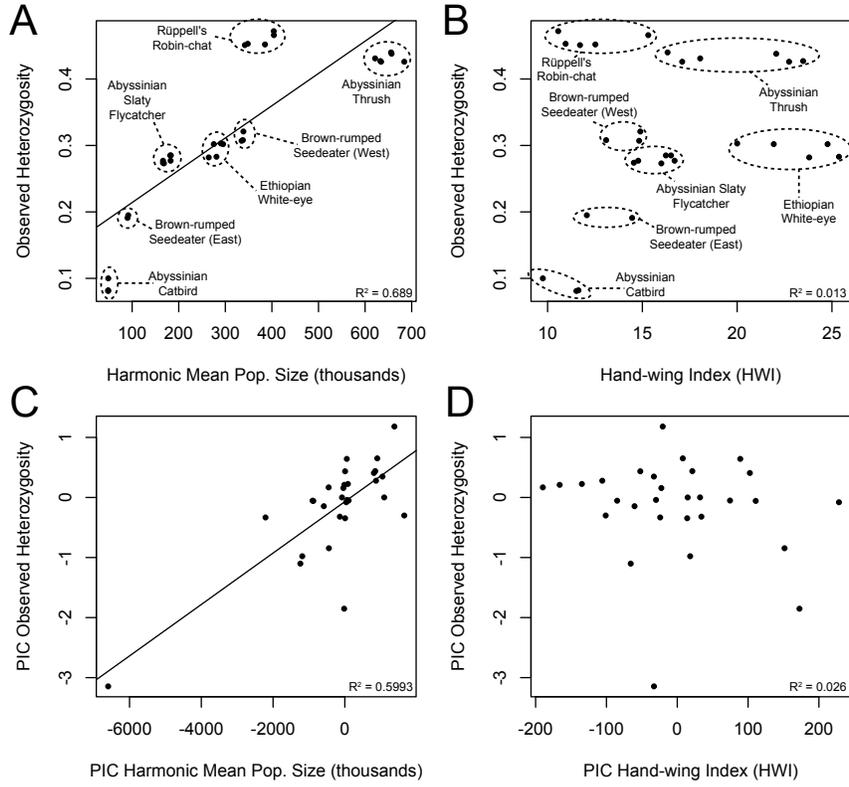
- Nadachowska-Brzyska, K., Li, C., Smeds, L., Zhang, G., & Ellegren, H. (2015). Temporal dynamics of avian populations during Pleistocene revealed by whole-genome sequences. *Current Biology*, *25* (10), 1375-1380.
- Naka, L. N., & Brumfield, R. T. (2018). The dual role of Amazonian rivers in the generation and maintenance of avian diversity. *Science Advances*, *4* (8), eaar8575.
- Nam, K., Mugal, C., Nabholz, B., Schielzeth, H., Wolf, J. B., Backstrom, N., . . . Ponting, C. P. (2010). Molecular evolution of genes in avian genomes. *Genome Biology*, *11* (6), R68.
- O'Connell, K. A., Hamidy, A., Kurniawan, N., Smith, E. N., & Fujita, M. K. (2018). Synchronous diversification of parachuting frogs (Genus *Rhacophorus*) on Sumatra and Java. *Molecular Phylogenetics and Evolution*, *123* , 101-112.
- Oliveros, C. H., Field, D. J., Ksepka, D. T., Barker, F. K., Aleixo, A., Andersen, M. J., . . . Braun, M. J. (2019). Earth history and the passerine superradiation. *Proceedings of the National Academy of Sciences, USA*, *116* (16), 7916-7925.
- Oswald, J. A., Overcast, I., Mauck III, W. M., Andersen, M. J., & Smith, B. T. (2017). Isolation with asymmetric gene flow during the nonsynchronous divergence of dry forest birds. *Molecular Ecology*, *26* (5), 1386-1400.
- Pages, H., Aboyoun, P., Gentleman, R., & DebRoy, S. (2017). Biostrings: Efficient manipulation of biological strings. *R Package Version 2.0* .
- Papadopoulou, A., Anastasiou, I., Keskin, B., & Vogler, A. P. (2009). Comparative phylogeography of tenebrionid beetles in the Aegean archipelago: the effect of dispersal ability and habitat preference. *Molecular Ecology*, *18* (11), 2503-2517.
- Paradis, E., Claude, J., & Strimmer, K. (2004). APE: analyses of phylogenetics and evolution in R language. *Bioinformatics*, *20* (2), 289-290.
- Peterman, W. E., Anderson, T. L., Ousterhout, B. H., Drake, D. L., Semlitsch, R. D., & Eggert, L. S. (2015). Differential dispersal shapes population structure and patterns of genetic differentiation in two sympatric pond breeding salamanders. *Conservation Genetics*, *16* (1), 59-69.
- Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics*, *155* (2), 945-959.
- Quinlan, A. R., & Hall, I. M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics*, *26* (6), 841-842.
- Reich, D., Thangaraj, K., Patterson, N., Price, A. L., & Singh, L. (2009). Reconstructing Indian population history. *Nature*, *461* (7263), 489.
- Reyes-Velasco, J., Manthey, J. D., Bourgeois, Y., Freilich, X., & Boissinot, S. (2018). Revisiting the phylogeography, demography and taxonomy of the frog genus *Ptychadena* in the Ethiopian highlands with the use of genome-wide SNP data. *PLoS One*, *13* (2), e0190440.
- Reyes-Velasco, J., Manthey, J. D., Freilich, X., & Boissinot, S. (2018). Diversification of African tree frogs (genus *Leptopelis*) in the highlands of Ethiopia. *Molecular Ecology*, *27* (9), 2256-2270.
- Romiguier, J., Gayral, P., Ballenghien, M., Bernard, A., Cahais, V., Chenuil, A., . . . Faivre, N. (2014). Comparative population genomics in animals uncovers the determinants of genetic diversity. *Nature*, *515* (7526), 261.
- Sayyari, E., & Mirarab, S. (2016). Fast coalescent-based computation of local branch support from quartet frequencies. *Molecular Biology and Evolution*, *33* (7), 1654-1668.
- Schiffels, S., & Durbin, R. (2014). Inferring human population size and separation history from multiple genome sequences. *Nature Genetics*, *46* (8), 919.

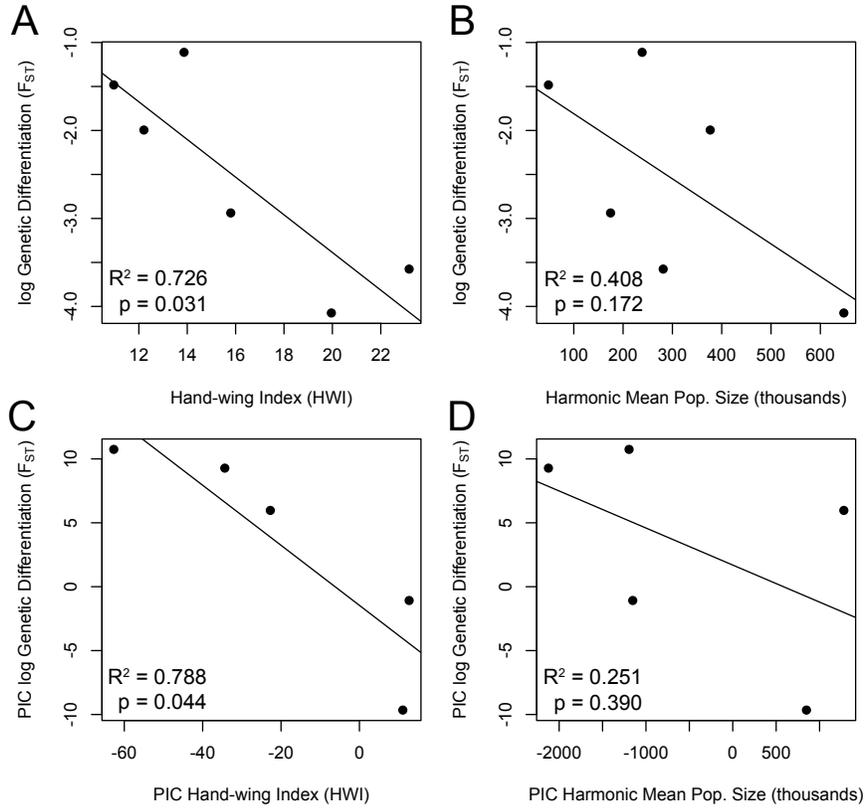
- Sepulchre, P., Ramstein, G., Fluteau, F., Schuster, M., Tiercelin, J.-J., & Brunet, M. (2006). Tectonic uplift and Eastern Africa aridification. *Science*, *313* (5792), 1419-1423.
- Shafer, A. B., Cullingham, C. I., Cote, S. D., & Coltman, D. W. (2010). Of glaciers and refugia: a decade of study sheds new light on the phylogeography of northwestern North America. *Molecular Ecology*, *19* (21), 4589-4621.
- Silvestrini, M., Junqueira, M. G., Favarin, A. C., Guerreiro-Filho, O., Maluf, M. P., Silvarolla, M. B., & Colombo, C. A. (2007). Genetic diversity and structure of Ethiopian, Yemen and Brazilian *Coffea arabica* L. accessions using microsatellites markers. *Genetic Resources and Crop Evolution*, *54* (6), 1367-1379.
- Smith, B. T., McCormack, J. E., Cuervo, A. M., Hickerson, M. J., Aleixo, A., Cadena, C. D., . . . Harvey, M. G. (2014). The drivers of tropical speciation. *Nature*, *515* (7527), 406.
- Stamatakis, A. (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*, *30* (9), 1312-1313.
- Sukumaran, J., & Holder, M. T. (2010). DendroPy: a Python library for phylogenetic computing. *Bioinformatics*, *26* (12), 1569-1571.
- Tacutu, R., Thornton, D., Johnson, E., Budovsky, A., Barardo, D., Craig, T., . . . Wang, J. (2017). Human Ageing Genomic Resources: new and updated databases. *Nucleic Acids Research*, *46* (D1), D1083-D1090.
- Taylor, B. (2020). Abyssinian Slaty-Flycatcher (*Melaenornis chocolatinus*), version 1.0. In Birds of the World (J. del Hoyo, A. Elliott, J. Sargatal, D. A. Christie, and E. de Juana, Editors). Cornell Lab of Ornithology, Ithaca, NY, USA. <https://doi.org/10.2173/bow.abysflf1.01>.
- Taylor, P. J., Lavrenchenko, L. A., Carleton, M. D., Verheyen, E., & Oosthuizen, C. J. (2011). Specific limits and emerging diversity patterns in east African populations of laminate-toothed rats, genus *Otomys* (Muridae: Murinae: Otomyini): revision of the *Otomys typus* complex. *Zootaxa*, *3024* , 1-66.
- van Balen, B., del Hoyo, J., Collar, N., & Kirwan, G. (2020). Heuglin's White-eye (*Zosterops poliogastrus*), version 1.0. In Birds of the World (S. M. Billerman, B. K. Keeney, P. G. Rodewald, and T. S. Schulenberg, Editors). Cornell Lab of Ornithology, Ithaca, NY, USA. <https://doi.org/10.2173/bow.heuwhe2.01>.
- Wang, K., Mathieson, I., O'Connell, J., & Schiffels, S. (2020). Tracking human population structure through time from whole genome sequences. *PLoS Genetics*, *16* (3), e1008552.
- Warren, W. C., Clayton, D. F., Ellegren, H., Arnold, A. P., Hillier, L. W., Kunstner, A., . . . Fairley, S. (2010). The genome of a songbird. *Nature*, *464* (7289), 757.
- Weeks, B. C., & Claramunt, S. (2014). Dispersal has inhibited avian diversification in Australasian archipelagoes. *Proceedings of the Royal Society of London, Series B: Biological Sciences*, *281* (1791), 20141257.
- Willing, E.-M., Dreyer, C., & Van Oosterhout, C. (2012). Estimates of genetic differentiation measured by FST do not necessarily require large sample sizes when using many SNP markers. *PLoS One*, *7* (8), e42649.
- Yalden, D., & Largen, M. (1992). The endemic mammals of Ethiopia. *Mammal Review*, *22* (3-4), 115-150.
- Zachos, F. E., & Habel, J. C. (2011). *Biodiversity hotspots: distribution and protection of conservation priority areas* : Springer Science & Business Media.
- Zelege, G., & Hurni, H. (2001). Implications of land use and land cover dynamics for mountain resource degradation in the Northwestern Ethiopian highlands. *Mountain research and development*, *21* (2), 184-192.
- Zhang, C., Rabiee, M., Sayyari, E., & Mirarab, S. (2018). ASTRAL-III: polynomial time species tree reconstruction from partially resolved gene trees. *BMC Bioinformatics*, *19* (6), 153.

Hosted file

Figure1_map.pdf available at <https://authorea.com/users/380686/articles/549339-varied-diversification-patterns-and-distinct-demographic-trajectories-in-ethiopian-montane-forest-bird-aves-passeriformes-populations-separated-by-the-great-rift-valley>







Hosted file

ethiopia_grv_table1.xlsx available at <https://authorea.com/users/380686/articles/549339-varied-diversification-patterns-and-distinct-demographic-trajectories-in-ethiopian-montane-forest-bird-aves-passeriformes-populations-separated-by-the-great-rift-valley>